

AD-A088 016

STANFORD UNIV CA DEPT OF OPERATIONS RESEARCH F/G 12/1
INDIVIDUAL VERSUS SOCIAL OPTIMIZATION IN COMMUTER ROUTE CHOICE.(U)
JAN 80 C E BELL N00014-76-C-0418

UNCLASSIFIED

TR-89

NL

1 of 1

34

1

END

DATA

FILED

4-80

DTIC



10

INDIVIDUAL VERSUS SOCIAL OPTIMIZATION IN COMMUTER ROUTE CHOICE,

BY

10 COLIN E. BELL

7 TECHNICAL REPORT, NO. 89

11 JANUARY 1980

14 TR-55

15 PREPARED UNDER CONTRACT

NSF-ENG 75-14847 (NR-047-061)

FOR THE OFFICE OF NAVAL RESEARCH

Frederick S. Hillier, Project Director

Reproduction in Whole or in Part is Permitted
for any purpose of the United States Government

This document has been approved for public release
and sale; its distribution is unlimited.

This research was supported in part by National Science Foundation
Grant ENG 75-14847 Department of Operations Research, Stanford University
issued as Technical Report No. 55.

DEPARTMENT OF OPERATIONS RESEARCH

STANFORD UNIVERSITY

STANFORD, CALIFORNIA

41276

Abstract

Each of N customers must select one of K commuter routes daily. Each commuter route is represented as a series of $M/D/1$ queuing stations with fixed travel times between stations. In the basic model, all customers arrive at the start of their selected route simultaneously; results which are optimal for the basic model are shown to be nearly optimal under a variety of more realistic arrival assumptions. In the basic model, the total wait on route k on any day when n_k customers use it is shown to be $\alpha_k n_k + \beta_k n_k^2$.

For the basic model three solutions are found and compared: (1) the deterministic social optimum, an assignment of individual customers to routes minimizing total wait per day, (2) the best common mixed strategy minimizing average total wait per day when customers' individual average waits must all be equal and (3) the equilibrium common mixed strategy that would be used by every customer acting in self-interest. The three solutions differ systematically. Individuals using (3) use routes k with low $\alpha_k + \beta_k$ with higher frequency than in (2) or (1). Since $\alpha_k + \beta_k$ is the total wait on route k if only one customer uses it, self-interested individuals overload those routes which would be fastest in the absence of congestion.

Accession For	
NTIS GRA&I	
DOC TAB	
Unannounced	
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or special
A	

Section 1 introduces the model assumptions. In Section 2 a deterministic social optimum is found. Section 3 explores mixed strategies and compares them to the deterministic social optimum. Best mixed strategies within certain classes can be found by solving revised deterministic social optimum problems by the methodology of Section 2. Section 4 examines more general arrival assumptions.

1. Assumptions

Many authors (e.g. in [1] and references 1, 3, 8-11, of [1]) have shown that customers acting in their self-interest in deciding whether or not to enter a queuing facility tend to cause more congestion at the facility than would be socially desirable. This is a result of an individual's failure to consider the inconvenience caused to later arrivals when he/she enters. In all of the above models customers decide whether or not to enter a queuing system based on a comparison of their perceived service reward and cost of waiting. In our model, customers will not have a choice of entering or not; they must enter the system but will have a choice of server.

Choosing a daily commuter route is perhaps the most appropriate example of our model. Assumptions are:

- (1) There is a finite calling population of known size, N .
- (2) There are K service facilities, $(1, 2, \dots, K)$.
- (3) Each customer is required to choose a service facility (e.g., commuter route) daily.
- (4) The arrival pattern of customers and the resulting congestion is concentrated at one period of the day. Steady state behavior

which applies to any time of day cannot be analyzed. However, the long run tendencies of customers to make certain choices of facility from day to day can be studied.

- (5) If n_k customers choose facility k on any day, the total number of customer hours spent in facility k is given by the simple expression

$$\alpha_k n_k + \beta_k n_k^2 \quad (1)$$

where α_k, β_k are known positive constants.

- (6) Customers have knowledge of assumptions (1)-(5) but must choose a service facility without knowledge of any other customers' choices.

In the context of choosing a commuter route, the following simple model gives rise to a total delay of the form $\alpha_k n_k + \beta_k n_k^2$ when n_k customers select route k . Although highly idealized, this model may provide some justification for using an expression as simple as $\alpha_k n_k + \beta_k n_k^2$ to approximate the total delay on a real commuter route.

We view route k as a system of single server queuing stations in series with constant service times and unlimited queue capacity. Regardless of the congestion level, there are also known constant travel times between stations. Along route k servers $1, 2, \dots, c_k$ provide services of constant duration $\alpha_{k1}, \alpha_{k2}, \dots, \alpha_{kc_k}$ respectively. Constant travel times t_{k1} and $\{t_{ki}, i=2, \dots, c_k\}$ are required to travel from the start of route k to server 1 and from server $i-1$ to server i ($i=2, \dots, c_k$). Figure 1 illustrates these assumptions. In addition, if n_k customers

select route k , it is assumed that all n_k customers arrive at the start of the route at the same instant.

Figure 1: Approximate Model of a Commuter Route

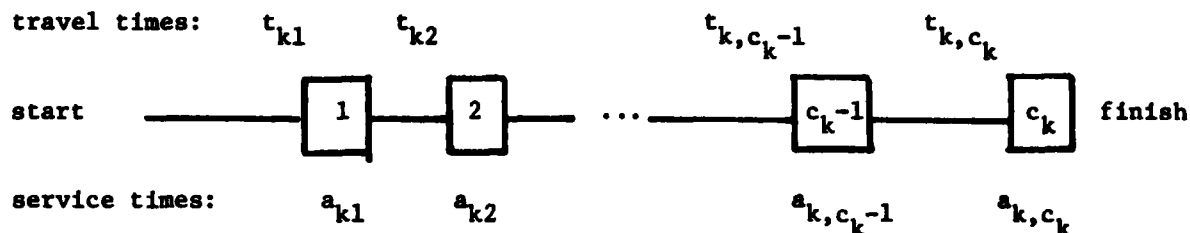
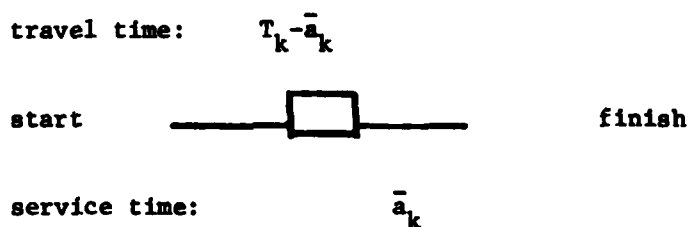


Figure 2: Equivalent Model



Generalizations of this last, rather restrictive assumption are discussed in Section 4. Our solutions will remain nearly optimal provided that arrivals occur at a reasonably fast rate. Precise conditions are given in Section 4.

The model of route k in Figure 1 can be further simplified. Let $\bar{a}_k = \max_{1 \leq i \leq c_k} \{a_{ki}\}$; \bar{a}_k is the service duration of the slowest (or bottleneck) server on route k . Let $T_k = \sum_{i=1}^{c_k} (t_{k,i} + a_{k,i})$; T_k is the total time to pass through route k for a customer who encounters no queuing delays at any server. If n_k customers enter route k at time 0, they

leave route k \bar{a}_k time units apart at times $T_k, T_k + \bar{a}_k, T_k + 2\bar{a}_k, \dots, T_k + (n_k - 1)\bar{a}_k$. This departure pattern results from the fact that if n_k customers arrive at the i -th server at times $t, t+a, t+2a, \dots, t+(n_k-1)a$, then they arrive at the $i+1$ st server at times $t+a_{k,i} + t_{k,i+1}, t+a_{k,i} + t_{k,i+1} + \max\{a, a_{k,i}\}, t+a_{k,i} + t_{k,i+1} + 2 \max\{a, a_{k,i}\}, \dots, t+a_{k,i} + t_{k,i+1} + (n_k-1) \max\{a, a_{k,i}\}$.

The same sequence of departure times $T_k, T_k + \bar{a}_k, T_k + 2\bar{a}_k, \dots, T_k + (n_k - 1)\bar{a}_k$ would result from the equivalent model shown in Figure 2. The total delay is $n_k T_k + 0.5(n_k - 1)n_k \bar{a}_k$, an expression of the form (1).

In Figure 2, the average total delay is unchanged if we assume that the server provides independent services of random duration with expected value \bar{a}_k . We henceforth assume that route k can be represented by a single server with average service duration \bar{a}_k and a fixed travel time $T_k - \bar{a}_k$ leading up to the server. The expected total delay encountered by n_k customers is then $n_k(T_k - \bar{a}_k) + 0.5n_k(n_k + 1)\bar{a}_k = n_k T_k + 0.5(n_k - 1)n_k \bar{a}_k$, an expression of the form (1) with

$$\alpha_k = T_k - \bar{a}_k / 2, \quad \beta_k = \bar{a}_k / 2 \quad (2)$$

Three solutions will be found and compared. In Section 2 the deterministic social optimum problem (DSOP) is solved. This solution assigns individual customers to routes to minimize the expected total wait for all customers. It is discriminatory in the sense that customers assigned to route k may have longer or shorter average waits than customers assigned to route k' . Customers acting in their self-interest could not be expected to follow such an assignment. Such a social optimum could

be attained only if customers were forced by a dictator to adhere to their assigned routes or if a set of incentives (tolls and rebates) could be provided which would convince any customer that it would be unwise to deviate from his assignment. Although such methods of enforcing the DSOP solution may be unrealistic, the DSOP solution is important because: (1) other solutions will be found by solving revised DSOP problems by the same methodology, (2) the objective function value provides a lower bound on the overall average wait, a benchmark for comparison with other solutions and (3) the DSOP solution allows a behavioral comparison with other solutions. When given free choice, individuals cause excess congestion in certain routes, less in others in comparison to the DSOP solution; roughly speaking, the routes which become over-congested are those with small T_k values, i.e., the routes which would be fastest at (off-peak travel) times when there are no delays due to queuing. These behavioral differences will be stated more precisely below.

The second solution (found in Section 3) is the common mixed strategy which minimizes expected total wait over all common mixed strategies. A mixed strategy (x_1, x_2, \dots, x_K) , $\sum_{k=1}^K x_k = 1$, $x_k \geq 0$, is a probability distribution over routes. By a common mixed strategy (x_1, x_2, \dots, x_K) we assume that every customer chooses a route on each day by drawing from this probability distribution. Since each customer has exactly the same available information at the time he makes a decision, it seems reasonable that all customers should adopt a common strategy. The numbers of customers (N_1, N_2, \dots, N_K) using routes 1, 2, ..., K on any day has a multinomial distribution with parameters N, x_1, x_2, \dots, x_K

representing N independent trials with K possible outcomes where outcome k has probability x_k on any trial. The best common mixed strategy (BCMS) solution has one advantage over the DSOP solution: the average wait is the same for every customer on any day. Thus this solution is more equitable. However, it is typically not an equilibrium solution; any individual acting in self-interest would be motivated to use a different strategy if he knew that all others would stick to (x_1, x_2, \dots, x_K) . In the absence of an incentive mechanism or a dictator, the instability of the BCMS solution makes its objective function value (overall expected wait) impossible to attain by N individuals acting in self-interest.

The third solution (found in Section 3) is the common mixed strategy equilibrium (CMSE). This is the common mixed strategy that would be used by self-interested customers without any tolls, rebates or other incentives. An equilibrium strategy has the property that no single customer would be motivated to deviate from it if he knew that all other customers would adhere to it. The average overall wait is naturally higher with this solution than with DSOP, or BCMS solutions.

Table 1 gives four numerical examples illustrating differences between DSOP, BCMS and CMSE solutions. In these examples with $K=2$, $N=10,000$, the average number of customers using route 1 varies substantially. The average wait per customer varies to a lesser extent but can be as much as 8.25 percent higher in the CMSE solution when compared with the DSOP. Notice that the average numbers of customers using route 1 in the DSOP and BCMS solutions are almost identical and the fact that the actual number of customers using route 1 is random in the BCMS case adds little to the average wait per customer.

In Section 2 the deterministic social optimum problem (DSOP) is solved; in Section 3 mixed strategy solutions (BCMS and CMSE) are found and all 3 solutions compared. Our main result is that self-interested individuals reach an equilibrium (CMSE solution) which tends to overload routes k with sufficiently high values of T_k when compared to the BCMS solution. Thus self-interested customers make too much use during rush hour of the routes which would be fastest when no one else is on the road. Compared to the DSOP solution, self-interested customers overload routes k with sufficiently high values of $T_k + (N-1)\bar{a}_k / (2N+2)$.

2. Deterministic Social Optimum Problem (DSOP)

In the analysis below we ignore the fact that n_1, n_2, \dots, n_K must all be integers. Under social optimization n_1, n_2, \dots, n_K must be set to minimize total delay in all systems combined. The DSOP is then the nonlinear programming problem:

$$\begin{aligned} &\text{minimize} && \sum_{k=1}^K (\alpha_k n_k + \beta_k n_k^2) \\ &\text{subject to} && \sum_{k=1}^K n_k = N \end{aligned} \quad (3)$$

$$n_k \geq 0, \quad k = 1, 2, \dots, K$$

The optimal solution $(n_1^*, n_2^*, \dots, n_K^*)$ is easily found by writing a Laggrangean

$$Z(n_1, n_2, \dots, n_K, \lambda) = \sum_{k=1}^K (\alpha_k n_k + \beta_k n_k^2) - \lambda \left(\sum_{k=1}^K n_k - N \right)$$

and evaluating partial derivatives

$$\frac{\partial Z(n_1, n_2, \dots, n_k, \lambda)}{\partial n_k} = \alpha_k + 2\beta_k n_k - \lambda. \quad (4)$$

$(n_1^*, n_2^*, \dots, n_k^*)$ must satisfy

$$\alpha_k + 2\beta_k n_k^* - \lambda^* = 0 \quad \text{if } 0 < n_k^* \leq N \quad (5)$$

$$\alpha_k - \lambda^* \geq 0 \quad \text{if } n_k^* = 0. \quad (6)$$

From (5) and (6) $n_k^* = [\lambda^* - \alpha_k]^+ / 2\beta_k$ where $a^+ = \max\{a, 0\}$. If $n_k(\lambda) \equiv [\lambda - \alpha_k]^+ / 2\beta_k$, then λ^* is the unique value of λ where $\sum_{k=1}^K n_k(\lambda) = N$. But $\sum_{k=1}^K n_k(\lambda)$ is a piecewise linear convex non-decreasing function of λ with kinks at $\alpha_1, \alpha_2, \dots, \alpha_K$. Thus λ^* (and hence $n_1^*, n_2^*, \dots, n_K^*$) can be found by a simple search.

3. Mixed Strategy Solutions

Under the DSOP solution $(n_1^*, n_2^*, \dots, n_K^*)$ the total wait in facility k is $\alpha_k n_k^* + \beta_k n_k^{*2}$; thus the average wait per customer in facility k is $\alpha_k + \beta_k n_k^*$. Thus a customer would be tempted to switch from facility k to facility ℓ if $\alpha_k + \beta_k n_k^* > \alpha_\ell + \beta_\ell (n_\ell^* + 1)$. Since (5) implies $\alpha_k + 2\beta_k n_k^* = \alpha_\ell + 2\beta_\ell n_\ell^*$, a customer would be tempted to switch from k to ℓ if $\alpha_k + \beta_k n_k^* + z(\alpha_k + 2\beta_k n_k^*) > \alpha_\ell + \beta_\ell (n_\ell^* + 1) + z(\alpha_\ell + 2\beta_\ell n_\ell^*)$ where z is an arbitrary

constant. Setting $z = -1$ and multiplying both sides by -1 yields the equivalent condition: $\beta_k n_k^* < \beta_\ell (n_\ell^* - 1)$. Setting $z = -0.5$ and multiplying both sides by 2 yields an alternative equivalent condition: $\alpha_k > \alpha_\ell + 2\beta_\ell$.

There are typically one or more equilibria involving different mixed strategies for different customers. However, since every customer has the same information, we concentrate on the case where all customers use the same mixed strategy. First, we can find the BCMS, the common mixed strategy which minimizes average wait for all customers. If every customer uses this strategy, an equilibrium generally does not result. The BCMS generally differs from the strategy $(n_1^*/N, n_2^*/N, \dots, n_K^*/N)$ which one might consider based on the DSOP allocation $(n_1^*, n_2^*, \dots, n_K^*)$. If each customer uses the mixed strategy (x_1, x_2, \dots, x_K) , then the actual number of customers selecting routes $1, 2, \dots, K$ is a multinomial random vector (X_1, X_2, \dots, X_K) with $E(X_k) = Nx_k$, $\text{Var}(X_k) = Nx_k(1-x_k)$ and $E(X_k^2) = Nx_k + N(N-1)x_k^2$. We then wish to find (x_1, x_2, \dots, x_K) to minimize

$$E\left\{\sum_{k=1}^K \alpha_k X_k + \beta_k X_k^2\right\} = \sum_{k=1}^K \{\alpha_k E(X_k) + \beta_k E(X_k^2)\}.$$

Substituting yields

$$\sum_{k=1}^K \{\alpha_k Nx_k + \beta_k [Nx_k + N(N-1)x_k^2]\} = \sum_{k=1}^K \{(\alpha_k + \beta_k)Nx_k + N(N-1)x_k^2\}.$$

Letting $n_k = Nx_k$ gives

$$\sum_{k=1}^K \{(\alpha_k + \beta_k)n_k + (N-1)n_k^2\beta_k/N\}.$$

This can be solved by finding the DSOP solution $(n_1', n_2', \dots, n_K')$ when

$$\alpha_k' = \alpha_k + \beta_k, \quad \beta_k' = (N-1)\beta_k/N \quad (7)$$

and then finding $(x_1', x_2', \dots, x_K') = (n_1'/N, n_2'/N, \dots, n_K'/N)$.

To find the common mixed strategy equilibrium $(\hat{x}_1, \hat{x}_2, \dots, \hat{x}_K)$ assume that customers 2 through N use this strategy. The number of these customers selecting routes 1, 2, ..., K respectively on any day is thus a multinomial random vector (N_1, N_2, \dots, N_K) with $E(N_k) = (N-1)\hat{x}_k$. By using route k with n_k others present, customer 1 can attain a conditional average wait of $T_k - \bar{a}_k + (1 + 5n_k)\bar{a}_k = T_k + 5n_k\bar{a}_k$. Since $E(N_k) = (N-1)\hat{x}_k$, customer 1 can attain an average wait of $T_k + 5(N-1)\hat{x}_k\bar{a}_k$ by selecting route k. Using (2)

$$T_k + 5(N-1)\hat{x}_k\bar{a}_k = \alpha_k + \beta_k + (N-1)\hat{x}_k\beta_k = \alpha_k + [1 + (N-1)\hat{x}_k]\beta_k \quad (8)$$

For an equilibrium we must have equal average waits on every route used (i.e. with $\hat{x}_k > 0$). Thus

$$\alpha_k + [1 + (N-1)\hat{x}_k]\beta_k - \hat{\lambda} = 0 \quad \text{if } 0 < \hat{x}_k < 1 \quad (9)$$

If there are unused routes ($\hat{x}_k = 0$), then no positive value of \hat{x}_k can attain an average wait of $\hat{\lambda}$.

Thus

$$\alpha_k + \beta_k - \hat{\lambda} \geq 0 \quad \text{if } \hat{x}_k = 0 \quad (10)$$

Conditions (8)-(9) are analogous to conditions (5)-(6). In fact, this equilibrium can be found by solving a DSOP with

$$\hat{\alpha}_k = \alpha_k + \beta_k, \quad \hat{\beta}_k = (N-1)\beta_k/2N \quad (11)$$

Denoting the solution to this DSOP by $(\hat{n}_1, \hat{n}_2, \dots, \hat{n}_K)$ we then have $(\hat{x}_1, \hat{x}_2, \dots, \hat{x}_K) = (\hat{n}_1/N, \hat{n}_2/N, \dots, \hat{n}_K/N)$.

The mixed strategy solutions BCMS and CMSE can be found by solving revised DSOP's with parameters given by (7) and (11) respectively. To compare DSOP solutions to different problems, recall that the original DSOP solution $(n_1^*, n_2^*, \dots, n_K^*)$ satisfies (5) and (6) and thus

$$n_k^* = [\lambda^* - \alpha_k]^+ / 2\beta_k \quad (12)$$

where λ^* is the unique value such that

$$\sum_{k=1}^K [\lambda^* - \alpha_k]^+ / 2\beta_k = N \quad (13)$$

From (12) and (13) it is easy to verify that a revised DSOP with $\alpha_k' = c\alpha_k$, $\beta_k' = c\beta_k$, $c > 0$ has the same solution $(n_1^*, n_2^*, \dots, n_K^*)$ found from (12) and (13)

by replacing α_k , β_k and λ^* by $c\alpha_k$, $c\beta_k$ and $c\lambda^*$ respectively. We thus state Lemma 1 without proof.

Lemma 1: Two N customer DSOP's with parameters $\{\alpha_k, \beta_k, k=1, 2, \dots, K\}$ and $\{c\alpha_k, c\beta_k, k=1, 2, \dots, K\}$ respectively have the same optimal solution provided $c > 0$.

This device of rescaling a DSOP's parameters allows alternative representations of (7) and (11). For (7) with $c=N/(N-1)$ we get

$$\alpha_k' = N(\alpha_k + \beta_k)/(N-1), \quad \beta_k' = \beta_k \quad (7')$$

For (11) with $c=2N/(N-1)$ we get

$$\hat{\alpha}_k = 2N(\alpha_k + \beta_k)/(N-1), \quad \hat{\beta}_k = \beta_k \quad (11')$$

Using (7') and (11'), BCMS and CMSE solutions can be found by solving a DSOP with revised α_k' s but original β_k' s. Theorem 2 compares solutions of 2 DSOP's which differ in this manner.

Theorem 2: Let $(n_1', n_2', \dots, n_K')$ solve the revised DSOP with parameters $\alpha_1', \alpha_2', \dots, \alpha_K'$ and $\beta_1, \beta_2, \dots, \beta_K$. Then if $n_k' = [\lambda' - \alpha_k']^+ / 2\beta_k$

$$n_k' > n_k^* \iff \alpha_k' - \alpha_k < \lambda' - \lambda^* \text{ and } \alpha_k' < \lambda'. \quad (14)$$

Proof: From (12) and $n_k' = [\lambda' - \alpha_k']^+ / 2\beta_k$

$$n_k' - n_k^* = \{[\lambda' - \alpha_k']^+ - [\lambda^* - \alpha_k^*]^+\} / 2\beta_k$$

Thus $n_k' > n_k^* \iff [\lambda' - \alpha_k']^+ - [\lambda^* - \alpha_k^*]^+ > 0$

$$\iff \lambda' - \alpha_k' > \lambda^* - \alpha_k^* \text{ and } \lambda' - \alpha_k' > 0$$

$$\iff \alpha_k - \alpha_k' < \lambda' - \lambda^* \text{ and } \alpha_k' < \lambda'.$$

Comparing the original DSOP solution to the BCMS solution found using (7'), (14) becomes

$$n_k' > n_k^* \iff [N(\alpha_k + \beta_k) / (N-1)] - \alpha_k < \lambda' - \lambda^* \text{ and } N(\alpha_k + \beta_k) / (N-1) < \lambda'$$

$$\iff \alpha_k + N\beta_k < (\lambda' - \lambda^*)(N-1) \text{ and } \alpha_k + \beta_k < (N-1)\lambda' / N.$$

Thus the BCMS solution has a higher average number of customers on route k than the DSOP solution when $\alpha_k + N\beta_k$ is sufficiently small. From (2), $\alpha_k + N\beta_k = T_k + (N-1)\bar{a}_k / 2$.

Comparing the DSOP solution to the CMSE solution found using (11'), (14) becomes

$$\hat{n}_k' > n_k^* \iff 2N(\alpha_k + \beta_k) / (N-1) - \alpha_k < \hat{\lambda} - \lambda^* \text{ and } 2N(\alpha_k + \beta_k) / (N-1) < \hat{\lambda}$$

$$\iff (N+1)\alpha_k + 2N\beta_k < (N-1)(\hat{\lambda} - \lambda^*) \text{ and } \alpha_k + \beta_k < (N-1)\hat{\lambda} / 2N.$$

Thus the CMSE solution has a higher average number of customers on route k than the DSOP solution when $(N+1)\alpha_k + 2N\beta_k$ or equivalently $\alpha_k + 2N\beta_k/(N+1)$ is sufficiently small. From (2):

$$\alpha_k + 2N\beta_k/(N+1) = T_k + (N-1)\bar{a}_k/2(N+1). \quad (15)$$

Comparing the BCMS solution to the CMSE solution in the same manner yields

$$\hat{x}_k > x_k' \iff \hat{n}_k > n_k' \iff \alpha_k + \beta_k < \text{constant and } \hat{n}_k > 0.$$

From (2), $\alpha_k + \beta_k = T_k$, exactly the expected travel time for route k in the absence of congestion. Customers acting in their own interest have a tendency to overload routes which are fastest in the absence of congestion when compared to the BCMS solution. Since in most practical examples $T_k \gg \bar{a}_k$ routes with sufficiently small T_k are likely to also have sufficiently small $T_k + (N-1)\bar{a}_k/2(N+1)$. Thus from (15) the same routes will often be overloaded by individuals in comparison to the original DSOP solution; this fact is also suggested by the limited computational evidence in Table 1 where there is little difference between DSOP and BCMS solutions.

4. Modifying the Arrival Assumption

Although commuters typically wish to gain access to a route system at epochs which lie in a relatively short time interval, our assumption

that all customers arrive at the same instant is at best an approximation. First, we state sufficient conditions on modified arrival assumptions to guarantee that the DSOP solution under instantaneous arrival remains nearly optimal. Later, additional sufficient conditions will be stated for the mixed strategy problems. Throughout this section we view each of the K facilities as a single server model as in Figure 2; service times may be stochastic.

Let $(n_1^*, n_2^*, \dots, n_K^*)$ solve the original instantaneous arrival DSOP yielding optimal total wait $W = \sum_{k=1}^K \{(T_k - 0.5\bar{a}_k)n_k^* + 0.5\bar{a}_k n_k^{*2}\}$. Assuming that all customers arrive at time 0, let D_1, D_2, \dots, D_K be the (possibly random) times at which the last customer departs servers 1, 2, ..., K respectively. (If $n_k^* = 0$, then $D_k \equiv 0$.) Now consider a modified arrival process in which the first arrival occurs at time 0. Using $(n_1^*, n_2^*, \dots, n_K^*)$ and any particular sequential assignment of customers to servers, let D_1', D_2', \dots, D_K' be the times at which the last customer departs servers 1, 2, ..., K respectively. Then we can state and prove Theorem 3.

Theorem 3: If $\epsilon \equiv \sum_{k=1}^K n_k^* E\{D_k' - D_k\}$, then $(n_1^*, n_2^*, \dots, n_K^*)$ yields an expected total wait which is ϵ - optimal for the DSOP with modified arrival process.

Proof: Let E_1, E_2, \dots, E_N be the epochs at which customers 1, 2, ..., N are served using $(n_1^*, n_2^*, \dots, n_K^*)$ in the original problem. Then $W = E\{\sum_{i=1}^N E_i\}$. Let E_1', E_2', \dots, E_N' be the corresponding service completion

epochs and A_1, A_2, \dots, A_N the corresponding arrival epochs under the modified arrival process assumption. Then $W' \equiv$ expected total wait = $E\{\sum_{i=1}^N E_i'\} - E\{\sum_{i=1}^N A_i\}$. If customer 1 is assigned to route k, then

$$E\{E_1'\} - E\{E_1\} \leq E\{D_k'\} - E\{D_k\}. \quad \text{Thus } 0 \leq E\{\sum_{i=1}^N E_i'\} - E\{\sum_{i=1}^N E_i\} \leq \epsilon \text{ and hence}$$

$$W' \leq W + \epsilon - E\{\sum_{i=1}^N A_i\}.$$

Any other allocation $(\hat{n}_1, \hat{n}_2, \dots, \hat{n}_K)$ with expected total wait \hat{W} in the original problem must have expected total wait $\hat{W}' \geq \hat{W} - E\{\sum_{i=1}^N A_i\}$ in the modified problem. Thus $\hat{W}' \geq W' - \epsilon$.

Theorem 3 admits any arrival process which does not significantly inconvenience any server. For example, provided that we can guarantee that each server will work without an idle period between services of any two customers, ϵ becomes $\sum_{k=1}^K n_k^* E\{\bar{A}_k\}$ where \bar{A}_k is the time that the first customer is assigned to route k. ($\bar{A}_k \equiv 0$ if $n_k^* = 0$.) If, for example, customers arrive at times $(i-1)t$, $i=1, 2, \dots, N$ and service durations in every queue are constant and longer than Kt , then by assigning the first K customers to queues $1, 2, \dots, K$ in descending order of n_k , we can attain $\epsilon \leq N(K-1)t/2$. The average wait per customer is then within $(K-1)t/2$ of optimal.

An analogue of Theorem 3 holds for the mixed strategy cases. However, we must assume that a modified arrival process gives no customer any information that can be used to advantage. Although customers may be aware of the overall arrival process assumptions, they must be restricted to choose a mixed strategy which is independent of their eventual

to the original instantaneous arrival model will be very close to optimal provided conditions (1) and (2) directly above hold with very high probability.

In the equilibrium case the average wait per customer on route k is $E\{\sum_{i=1}^{N_k} E_i'/N_k\} - E\{\sum_{i=1}^{N_k} A_i/N_k\}$ where N_k is the random number of customers using route k and both quantities in braces are defined to be 0 when $N_k=0$. Comparing the average waits per customer on routes k and m where $x_k > 0$ and $x_m > 0$ (so that routes k and m are used) we first notice that $E\{\sum_{i=1}^{N_k} A_i/N_k\} = E\{\sum_{i=1}^{N_m} A_i/N_m\}$. Thus the difference between the average wait per customer on route k and that on route m is

$$\begin{aligned} & E\{\sum_{i=1}^{N_k} E_i'/N_k\} - E\{\sum_{i=1}^{N_m} E_i'/N_m\} \\ & \leq E\{\sum_{i=1}^{N_k} E_i'/N_k\} - E\{\sum_{i=1}^{N_m} E_i/N_m\} \\ & \leq E\{\sum_{i=1}^{N_k} E_i/N_k\} + E\{D_k' - D_k\} - E\{\sum_{i=1}^{N_m} E_i/N_m\} \\ & = E\{D_k' - D_k\} \end{aligned} \quad (16)$$

By reversing the roles of k and m we can also show that

$$E\{\sum_{i=1}^{N_k} E_i'/N_k\} - E\{\sum_{i=1}^{N_m} E_i'/N_m\} \geq -E\{D_m' - D_m\} \quad (17)$$

As long as $E\{D_k' - D_k\}$ and $E\{D_m' - D_m\}$ are small there is little difference in the average wait per customer between routes.

As an example of the magnitude of ϵ consider the case with $K=2$, $N=10^3$, $T_1=T_2=15$, exponential service with $\bar{a}_1=\bar{a}_2=.06$. By symmetry, the DSOP solution is (500,500) and both mixed strategy solutions are (.5,.5). In the absence of traffic the average commuter time is 15, in the DSOP solution it is 29.97. If interarrival times are exponential with expected

value 1.5×10^{-2} and all customers use the mixed strategy $(.5, .5)$, then up to time 30×10^{-2} the number of arrivals in each system is Poisson (10). Since the arrival rate in each system is 33.3 and the service rate 16.7, each server is almost certain to be busy from time $15 - .06/2 + 30 \times 10^{-2}$ until his system is emptied. Thus $\epsilon = 10^3 \times 30 \times 10^{-2} = 300$ is conservative and the average wait per customer is guaranteed by Theorem 3 to be within 0.3 of the optimal average wait, i.e. an error of less than one percent.

Reference

1. Lippman, S. A. and S. Stidham, Jr., "Individual Versus Social Optimization in Exponential Congestion Systems," Operations Research 25, pp. 233-247 (1977).

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DD FORM 1473
1 JAN 73

EDITION OF 1 NOV 68 IS OBSOLETE
S/N 0102-014-6601

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Technical Report No. 89

Abstract

Each of N customers must select one of K commuter routes daily. Each commuter route is represented as a series of $M/D/1$ queuing stations with fixed travel times between stations. In the basic model, all customers arrive at the start of their selected route simultaneously; results which are optimal for the basic model are shown to be nearly optimal under a variety of more realistic arrival assumptions. In the basic model, the total wait on route k on any day when n_k customers use it is shown to be $\alpha_k n_k + \beta_k n_k^2$.

For the basic model three solutions are found and compared: (1) the deterministic social optimum, an assignment of individual customers to routes minimizing total wait per day, (2) the best common mixed strategy minimizing average total wait per day when customers' individual average waits must all be equal and (3) the equilibrium common mixed strategy that would be used by every customer acting in self-interest. The three solutions differ systematically. Individuals using (3) use routes k with low $\alpha_k + \beta_k$ with higher frequency than in (2) or (1). Since $\alpha_k + \beta_k$ is the total wait on route k if only one customer uses it, self-interested individuals overload those routes which would be fastest in the absence of congestion.

**DATE
FILMED**

—8